

Contents

I	THEORETICAL FOUNDATIONS	1
1	Finite probability theory	3
1.1	Discrete random variables	4
1.2	Structures of conditional independence	6
1.3	Decomposability	12
1.4	Information-theoretic notions	16
1.5	Survey of symbols	22
2	Operator of composition	25
2.1	Basic properties	25
2.2	Anticipating operator	29
2.3	Projection	32
2.4	Iterative Proportional Fitting	33
3	Compositional models	37
3.1	Perfect models	38
3.2	Decomposable models	43
3.3	Marginalization	47
3.4	Conditioning	52
3.5	Causal models	57
4	Independence structure of models	63
4.1	Persegrams	64
4.2	Simple trails	65
4.3	Avoiding trails	67
5	Avoiding model overfitting	69
5.1	Information and complexity	70
5.2	Huffman code	72
5.3	Coding Data	74
5.4	Coding Models	78
5.5	Model Simplification	81

9.4.5 Convert to distribution 121

10 Others 123

10.1 Save and load 123

10.2 Referencing 123

Appendix: List of functions 129

6	Data mining example	85
II	SYSTEM MANUAL	95
7	Starting with MUDIM	97
7.1	Install R	97
7.2	Install MUDIM	98
8	Probability distribution	99
8.1	R object	99
8.1.1	Probability table	100
8.1.2	Names of random variables	104
8.1.3	Additional information	105
8.2	Manipulations with probability distributions	105
8.2.1	Marginal distribution	105
8.2.2	Product	106
8.2.3	Composition	107
8.2.4	Anticipating operator	108
8.3	Information-theoretic notions	108
8.3.1	Shannon entropy	108
8.3.2	Kullback-Leibler divergence	109
8.3.3	Mutual information	110
8.3.4	Conditional mutual information	110
8.3.5	Multi-information	110
8.3.6	Conditional multiinformation	111
9	Compositional model	113
9.1	R Object	113
9.2	Insert distribution	114
9.3	Model properties	115
9.3.1	Basic overview	115
9.3.2	Name and information	115
9.3.3	Length	116
9.3.4	Dimension	116
9.3.5	Structure	116
9.3.6	Random variables	116
9.3.7	Decomposability	117
9.4	Manipulations with model	117
9.4.1	Marginalization	117
9.4.2	Perfectization	119
9.4.3	Conditioning	119
9.4.4	Decomposibility	120